

# An Overview of Model Discrimination Techniques in Polymerization Processes

Noushin Majdabadifarahani, Thomas Duever\*, and Alexander Penlidis

Institute for Polymer Research, Department of Chemical Engineering,  
University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1, Canada

\*Ryerson University, 350 Victoria Street, Toronto, ON M5B 2K3

Often during a modelling study, there are several models which seem to provide a reasonable description of the process based on theory or existing experimental data trends. To model the process, a decision need to be made about which model is “best”. The objective of model discrimination techniques is to select the “best model” among rival candidate models. The methodology is very general and useful in any discipline in Science and Engineering, where alternative models exist, such as pharmacokinetic and reservoir engineering, but we will draw more examples from the polymerization area, given the nature of research done within IPR.

The models are of the form:

$$y_{i,n} = f_i(\underline{x}_n, \underline{\theta}_i) + \epsilon_{i,n} \quad i = 1, 2, \dots, m$$

The experimental point  $x_n$  should be chosen using the criterion:

$$\max \sum_{i=1}^{m-1} \sum_{j=i+1}^m (\hat{y}_i - \hat{y}_j)^2$$

The above criterion, which has become the basis or other criteria or modifications that followed, is one of the first (if not the first) discrimination attempts and is referred to as the Hunter-Reiner criterion. An improvement on the above equation can be made by considering the variance associated with individual predictions:

$$\max \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{(\hat{y}_i - \hat{y}_j)^2}{\text{var}(\hat{y}_i - \hat{y}_j)}$$

The first step in model discrimination is computer simulation of the process/system in question. Given some preliminary data from the process, a mathematical model can be developed, which can capture at least the basic trends. Subsequently, several model discrimination methods can be applied. These model discrimination methods, ranging from simple ones to more complicated ones, will be reviewed briefly. These methods include Box-Hill (1967), Exact Entropy (1970), Hsiang and Reilly (1971), Buzzi-Ferraris and Forzatti (1983), and Burke et al. (1997). Examples will be shown from their application to different cases, including copolymerization systems.

A function of entropy was introduced by Box-Hill (1967), as follows:

$$\zeta = - \sum_{i=1}^m Pr(M_i) \ln Pr(M_i)$$

Although the Hsiang and Reilly (1971) criterion is different from the exact entropy criterion (Reilly 1970), they both employ a Bayesian analysis. This starts by assigning prior model probabilities and these probabilities are updated after each experiment until  $Pr(M_{i,n})$  becomes sufficiently large.

$$\sum_{i=1}^m Pr(M_{i,0}) = 1$$

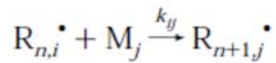
$$Pr(M_{i,n}) = Pr(M_{i,n-1}) L(M_i|y_n)$$

Buzzi-Ferraris and Forzatti (1983) developed a practical design criterion. They proposed that the  $n^{\text{th}}$  experiment be chosen to maximize:

$$T(x_n) = \frac{\sum_{i=1}^{m-1} \sum_{j=i+1}^m (\hat{y}_i - \hat{y}_j)^2}{(m-1)(ms^2 + \sum_{i=1}^m s_i^2)}$$

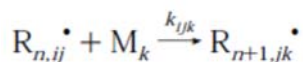
Once a model discrimination method is selected, one has to decide about available measurements (process responses) and related information on experimental error. For instance, which measurements are available and with what error? Copolymer composition, triad fractions, molecular weight averages, copolymerization rate, or combinations of the above? The success of such an analysis depends strongly on the information contained in the data.

An example of past work on copolymerization modeling is on the problem of discriminating between the terminal and penultimate models for copolymerization kinetics by Burke et al. (1997). The terminal model assumes that only the last unit on the growing chain influences monomer addition:



$$r_1 = \frac{k_{11}}{k_{12}}, \quad r_2 = \frac{k_{22}}{k_{21}}$$

The penultimate model assumes that the last two monomer units on the growing radical chain influence monomer addition:



$$r_{11} = \frac{k_{111}}{k_{112}}, \quad r_{21} = \frac{k_{211}}{k_{212}}$$

$$r_{22} = \frac{k_{222}}{k_{221}}, \quad r_{12} = \frac{k_{122}}{k_{121}}$$

Two radical reactivity ratios:

$$s_1 = \frac{k_{211}}{k_{111}}, \quad s_2 = \frac{k_{122}}{k_{222}}$$

If an attempt is made to fit the penultimate model to a system best described by the terminal model, joint confidence regions for the reactivity ratio estimates will include the points

$$\hat{r}_{11} = \hat{r}_{21}$$

$$\hat{r}_{22} = \hat{r}_{12}$$

$$\hat{s}_1 = \hat{s}_2 = 1$$

The work of Hill et al. (1982), on the styrene/acrylonitrile system, is typical of previous research on model discrimination. The copolymer compositions were measured and used to estimate terminal and penultimate reactivity ratios using instantaneous composition equations. Figure 1 is a plot of the data and the fitted curves.

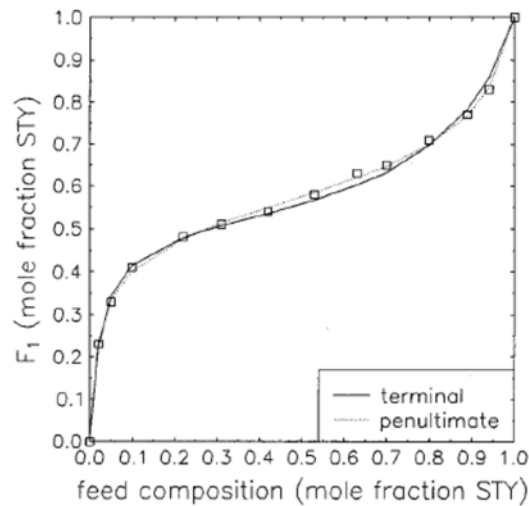


Fig. 1. Copolymer composition data for STY/AN from Hill et al. (1982)

These experiments do not seem to be optimal for either parameter estimation or model discrimination. Therefore, it is usually necessary to perform some initial parameter estimation experiments. The combined use of designed parameter estimation and statistical model discrimination experiments should improve our ability to discriminate between competing models.

Burke et al. (1997) classified statistical model discrimination experimental design criteria into maximum divergence-based and maximum entropy-based methods. These two categories seem to be able to cover all criteria which have been proposed in the literature. In an effort to discriminate between the terminal and penultimate model in copolymerization, Burke et al. (1997) found that triad fractions (i.e., monomer unit sequence information in the chain) provide good information for model discrimination. As with composition data, triad fractions depend only on monomer reactivity ratios. Table 1 and Table 2 show some results based on copolymer composition and triad fraction data.

Table 1. Results Based on Copolymer Composition Data

simulation model	model chosen as "best" at 95% confidence (number of simulation runs)		
	terminal	penultimate	neither
A. Buzzi-Ferraris and Forzatti Method			
terminal	7	4	16
strong penultimate	3	22	2
small penultimate	3	16	8
B. Exact Entropy Method			
terminal	19	6	2
strong penultimate	21	4	2
small penultimate	22	3	2
C. Hsiang and Reilly Method			
terminal	23	3	1
strong penultimate	17	8	2
small penultimate	25	2	0

Table 2. Results Based on Triad Fraction Data

simulation model	model chosen as "best" at 95% confidence (number of simulation runs)		
	terminal	penultimate	neither
A. Buzzi-Ferraris and Forzatti Method			
terminal	26	0	1
strong penultimate	0	27	0
small penultimate	0	27	0
B. Exact Entropy Method			
terminal	8	7	12
strong penultimate	0	27	0
small penultimate	0	27	0
C. Hsiang and Reilly Method			
terminal	17	9	1
strong penultimate	0	27	0
small penultimate	4	23	0

Figures 2 and 3 show the 95% joint confidence region for  $r_{21} = 6562$  and  $r_{11} = 0.5406$  after nine experiments, and the 95% joint confidence region for  $r_{22} = 4918$  and  $r_{12} = 0.7486$  based on the data of Maxwell et al. (1993), respectively. The cross represents the point estimates. It can be seen from Figure 3 that there is a great deal of uncertainty in the value of  $r_{12}$ .

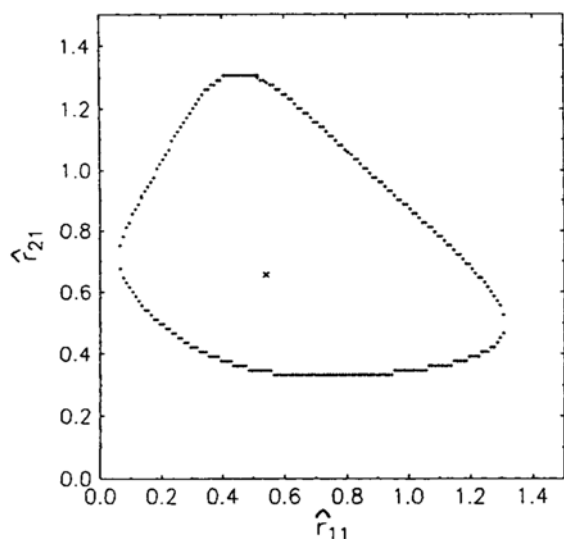


Fig 2. The 95% joint confidence region for  $r_{21}$  and  $r_{11}$

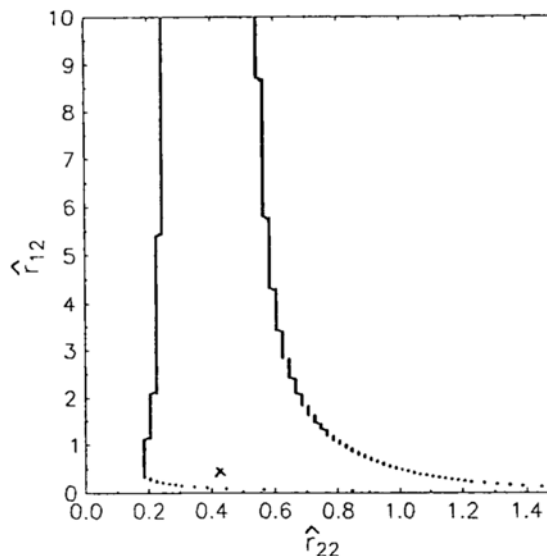


Fig 3. The 95% joint confidence region for  $r_{22}$  and  $r_{12}$

Landry et al. (2000) studied the application of weight average chain length for discrimination between the terminal and the penultimate copolymerization models. They found that combining weight average chain length with triad fractions yielded an even more powerful combination of measurements for model discrimination. In addition, Landry et al. (2000) investigated the effect of the presence of impurities during copolymerization on model discrimination, which may lead to incorrect choices about the copolymerization mechanism itself.

### Concluding remarks

An overview of the above methods will be given during the IPR conference along with a description of more advanced and recent efforts, like Sequential Bayesian Monte Carlo Model Discrimination (SBMCMD) approaches/strategies.

### References

- Box, G. E. P., Hill W. J., *Technometrics*, 1967, 9, 57
- Burke A. L., Duever T. A., and Penlidis A., *Ind. Eng. Chem. Res.* 1997, 36, 1016
- Burke A. L., Duever T. A, and Penlidis A., *J. of Pol. Sci.*, 1996, 34, 2665
- Buzzi-Ferraris G., and Forzatti P., *Chem. Eng. Sci.*, 1984, 39, 81
- Buzzi-Ferraris G., and Forzatti P., *Chem. Eng. Sci.* 1983, 38, 225
- Hill D. J. T., O'Donnell J. H., O'Sullivan, P. W., *Macromolecules*, 1982, 15, 960

- Hsiang T., and Reilly P. M., *Can. J. Chem. Eng.*, 1971, 49, 865
- Landry R., Penlidis A., and Duever T. A., *J. of Pol. Sci.*, 2000, 38, 2319
- Maxwell, I. A., Aerdt A. M., and German A. L., *Macromolecules*, 1993, 26, 1956
- Reilly, P. M., *Can. J. Chem. Eng.* 1970, 48, 168